

LECTURA RECOMENDADA

INTELIGENCIA ARTIFICIAL E INVESTIGACIÓN CIENTÍFICA

Inteligencia artificial e investigación científica.

- Introducción
- Uso de la inteligencia artificial en la investigación
- La importancia de los prompts
- Buenas prácticas.
- Uso ético de la inteligencia artificial

Índice rápido

| | |
|---|----|
| Introducción | 2 |
| Uso de la inteligencia artificial en la investigación | 3 |
| La importancia de los prompts..... | 9 |
| Uso ético de la inteligencia artificial generativa | 16 |

Nota de los formadores:

Este documento explora el impacto de la inteligencia artificial en la investigación científica, analizando su integración en diversas etapas del proceso investigador: desde la automatización de tareas y el análisis de grandes volúmenes de datos hasta la formulación de hipótesis y la comunicación de resultados. Se profundiza en el diseño de prompts efectivos como herramienta clave para optimizar la interacción con modelos de lenguaje, y se presentan buenas prácticas para su implementación. Finalmente, se examinan los desafíos éticos asociados al uso de la IA, con especial atención a la transparencia, la integridad académica y la regulación europea vigente.

Introducción

La irrupción de la inteligencia artificial (IA) ha transformado radicalmente múltiples aspectos de la vida cotidiana, desde la interacción con la tecnología hasta la organización del trabajo y los procesos de toma de decisiones. Aunque su visibilidad ha aumentado recientemente debido al auge de herramientas como ChatGPT, la IA se encuentra presente desde hace tiempo en actividades cotidianas que afectan directamente a las personas, como la navegación mediante GPS, el reconocimiento facial o las recomendaciones de compra en plataformas digitales. La IA constituye, en esencia, una disciplina orientada al desarrollo de sistemas capaces de llevar a cabo tareas complejas que, hasta hace poco, requerían la intervención de la inteligencia humana.

El funcionamiento de la IA se basa en diversas ramas tecnológicas fundamentales. Una de las más relevantes es el *machine learning* (ML) o aprendizaje automático, que permite a los sistemas aprender a partir del análisis de grandes volúmenes de datos, sin necesidad de una programación explícita para cada tarea. Dentro del ML, destaca el *deep learning* (DL) o aprendizaje profundo, técnica que se fundamenta en el uso de redes neuronales artificiales diseñadas para emular ciertos procesos cognitivos humanos. Estas redes posibilitan el reconocimiento de patrones complejos y la toma de decisiones con un grado creciente de precisión. Por su parte, el *natural language processing* (NLP) o procesamiento del lenguaje natural hace posible que los sistemas

interpreten y generen lenguaje humano, lo que ha permitido el desarrollo de aplicaciones como la traducción automática, la síntesis del habla o la visión por computador, empleada, por ejemplo, en la indexación automática de imágenes.

En los últimos años, se ha observado un notable desarrollo en el ámbito de la inteligencia artificial generativa (IAG), rama que combina aprendizaje profundo y procesamiento del lenguaje natural para la producción de contenido original en múltiples formatos: texto, imagen, sonido, voz o vídeo. Esta tecnología responde a instrucciones expresadas en lenguaje natural, lo que facilita su adopción en entornos no especializados. Un ejemplo destacado de esta innovación es la generación de lenguaje mediante modelos de lenguaje de gran tamaño (*large language models*, LLMs), capaces de redactar ensayos, responder a preguntas complejas, elaborar resúmenes, traducir textos o generar código de programación.

Entre las herramientas basadas en LLMs más reconocidas se encuentran ChatGPT (OpenAI), Gemini (Google), Claude (Anthropic) y Microsoft Copilot. Estas plataformas han democratizado el acceso a capacidades avanzadas de procesamiento lingüístico y producción de contenido, con aplicaciones extendidas en contextos educativos, profesionales y científicos. Su capacidad de adaptación a distintos dominios del conocimiento y su eficiencia en tareas que requieren interpretación semántica han contribuido a su consolidación en múltiples sectores.

Uno de los ámbitos donde la IA está demostrando un impacto especialmente significativo es el de la investigación científica. El uso de herramientas basadas en LLMs permite automatizar diversas fases del proceso de investigación, como la identificación de líneas de investigación, la [búsqueda](#) y análisis de información, el procesamiento de datos complejos, la selección de revistas para la publicación de resultados, o el formateado de referencias bibliográficas. Este tipo de soluciones contribuye a agilizar flujos de trabajo y permite centrar la atención en las tareas de mayor valor.

Uso de la inteligencia artificial en la investigación

El progreso de la IA ha ido acompañado de una integración progresiva en la investigación científica. Inicialmente, la IA se aplicó a problemas puntuales (como diagnósticos médicos o análisis estadístico), pero hoy forma parte del flujo de trabajo científico en múltiples disciplinas. Desde la automatización de tareas rutinarias de laboratorio hasta la generación de hipótesis originales, la IA está potenciando la velocidad y alcance de la investigación. La adopción de herramientas impulsadas por IA ha acelerado procesos tradicionalmente lentos o tediosos: por ejemplo, se dispone ya de asistentes digitales que ayudan a buscar bibliografía, analizar datos experimentales o traducir artículos académicos. De hecho, el uso extendido de la IA está revitalizando la investigación científica, ahorrando tiempo y optimizando recursos en la mayoría de sus

etapas.

Automatización de tareas rutinarias

Un primer ámbito de impacto de la IA es la automatización de tareas repetitivas o rutinarias en la labor científica. Muchas tareas que antes requerían la atención constante del personal investigador o técnico ahora pueden ser delegadas a sistemas inteligentes. Por ejemplo, en los laboratorios de biología molecular y química, existen robots controlados por IA que realizan procedimientos experimentales estándar (medición de reactivos, manejo de pipetas, cultivo de células, etc.) con mínima supervisión humana. Estos sistemas, guiados por algoritmos de aprendizaje automático, pueden controlar condiciones experimentales y recolectar datos de forma autónoma, liberando al personal científico de tareas mecánicas y permitiéndoles enfocarse en el análisis e interpretación más avanzado. Del mismo modo, en astronomía, la clasificación inicial de objetos celestes (como la identificación de galaxias en imágenes de telescopios) se ha automatizado con redes neuronales convolucionales capaces de reconocer patrones visuales; una tarea que solía involucrar a equipos de investigadores durante meses ahora puede completarse en horas con igual o mayor precisión. En ciencias sociales, otras tareas rutinarias, por ejemplo, la transcripción y codificación de entrevistas, la revisión de encuestas o la categorización de respuestas abiertas, se han visto facilitadas por herramientas de procesamiento de lenguaje natural, disminuyendo drásticamente el tiempo invertido en preparar los datos para su análisis. En humanidades digitales, la IA también apoya tareas como el reconocimiento óptico de caracteres en documentos históricos y su posterior etiquetado con metadatos,

agilizando la digitalización de archivos y la creación de bases de datos culturales. Estas aplicaciones muestran que la IA puede actuar como un asistente que mantiene en marcha procesos operativos de la investigación.

Además de acelerar el trabajo, la automatización inteligente contribuye a reducir errores humanos en procesos repetitivos. Los sistemas de visión por computadora (un área del aprendizaje profundo) pueden inspeccionar muestras o imágenes en busca de características específicas con una consistencia difícil de lograr manualmente. El resultado es una mayor eficiencia del método científico, donde las etapas más mecánicas se ejecutan con velocidad y exactitud, permitiendo a las y los investigadores concentrar su tiempo en el diseño de experimentos, la formulación de preguntas, la verificación de datos y la interpretación de resultados.

Análisis de grandes volúmenes de datos

La ciencia moderna se caracteriza por la generación de volúmenes masivos de datos, desde secuencias genómicas y registros de telescopios hasta datos de redes sociales y simulaciones físicas. La IA ha emergido como la herramienta indispensable para analizar grandes conjuntos de datos y extraer de ellos conclusiones significativas. A diferencia de los métodos estadísticos tradicionales, los algoritmos de IA pueden manejar cientos de variables simultáneamente y detectar estructuras complejas en los datos sin requerir modelos teóricos predefinidos. Esto resulta crucial en disciplinas como la física de partículas o la astrofísica, donde los experimentos producen cantidades de información abrumadoras.

Por ejemplo, los detectores del Gran Colisionador de Hadrones del CERN generan del orden de 60 terabytes de datos por segundo, aunque solo una fracción mínima de esos datos contiene eventos físicamente interesantes. Es inviable que el personal investigador examine manualmente tal flujo; en su lugar, algoritmos de aprendizaje automático en tiempo real filtran y reconocen patrones en las colisiones, identificando en milisegundos aquellas señales que podrían indicar nuevas partículas o fenómenos.

La capacidad de la IA para manejar *big data* también ha transformado la astronomía. En proyectos de búsqueda de exoplanetas, por ejemplo, se analizan los diminutos cambios de luminosidad de miles de estrellas simultáneamente para detectar tránsitos planetarios. Los algoritmos de aprendizaje automático pueden escudriñar estos enormes conjuntos de datos de manera rápida y precisa, detectando las sutiles disminuciones de brillo causadas por el paso de un planeta frente a su estrella. De forma análoga, en climatología se usan redes neuronales para integrar datos de satélites, estaciones meteorológicas y simulaciones oceánicas, identificando patrones en escalas espaciales y temporales que ayudan a mejorar los modelos de cambio climático. En ciencias sociales, la IA permite analizar millones de publicaciones en redes sociales o vastos corpus de texto en busca de tendencias culturales y comportamientos colectivos. Técnicas de procesamiento de lenguaje natural identifican temas recurrentes, sentimientos y conexiones semánticas en tiempo real, algo invaluable para estudios de opinión pública, epidemiología digital (seguimiento de brotes a través de buscadores) o análisis histórico de discursos en humanidades. En resumen, allí donde hay grandes volúmenes de datos, la IA actúa como un microscopio conceptual,

revelando regularidades y anomalías ocultas. Esto expande la capacidad humana de ver dentro de los datos: mientras que el ojo humano o las hojas de cálculo tradicionales se quedan cortos, los modelos de IA desentrañan la información latente, proporcionando una comprensión más profunda de sistemas complejos.

Apoyo en la revisión de literatura y formulación de hipótesis

Otra área de aplicación emergente de la IA en la ciencia es como asistente intelectual en la exploración bibliográfica y la generación de nuevas hipótesis. El crecimiento explosivo de la literatura científica hace prácticamente imposible que una persona abarque todo lo publicado incluso en un campo muy especializado. Mantenerse al día frente a la avalancha de información es un desafío enorme, pero las herramientas de IA están comenzando a brindar alivio. Existen motores de búsqueda académica impulsados por IA, como Underming, Consensus, Elicit o Scite, que pueden rastrear bases de datos de publicaciones y responder preguntas del usuario ofreciendo respuestas fundamentadas en la literatura existente. A diferencia de un buscador tradicional, estos sistemas analizan el contexto semántico de la pregunta y sintetizan hallazgos relevantes, incluso proporcionando referencias bibliográficas concretas. Por ejemplo, ante una pregunta compleja como: “¿Qué intervenciones psicológicas muestran mayor eficacia en el tratamiento de la ansiedad en adolescentes?”, un motor impulsado por inteligencia artificial puede analizar miles de artículos académicos y devolver un resumen estructurado con los enfoques más validados (como la terapia cognitivo-conductual o la terapia de aceptación y compromiso), indicando la magnitud del efecto, las poblaciones estudiadas y las fuentes principales. Además,

plataformas como SciSpace o Research Rabbit permiten visualizar redes de conocimiento, mostrando cómo se relacionan entre sí investigadores(as), enfoques teóricos y artículos clave. Esta capacidad de mapeo ayuda no solo a rastrear la evolución de un campo, sino también a descubrir patrones ocultos o áreas insuficientemente exploradas. Sin embargo, esta aparente exhaustividad puede ser engañosa. Aunque los sistemas de IA acceden a grandes volúmenes de información, su cobertura suele ser desigual, como un queso gruyere: algunos temas están muy bien representados (con múltiples estudios, revisiones sistemáticas y metaanálisis), mientras que otros quedan relegados por falta de datos, sesgo editorial o incluso por limitaciones idiomáticas y de acceso.

Los modelos de lenguaje avanzados como ChatGPT o aplicaciones como NotebookLM también se utilizan para resumir artículos y extraer conclusiones. Un investigador(a) puede proporcionar a un sistema de IA enlaces a artículos y pedir un resumen comparativo: la IA condensará metodologías y hallazgos principales, facilitando una revisión rápida. Además, la IA puede adoptar un rol más creativo al combinar información de fuentes diversas. Por ejemplo, algunos equipos de investigación emplean LLMs entrenados en literatura científica para proponer hipótesis o experimentos. La IA, puede generar preguntas inusuales o sugerir conexiones entre teorías disparesEn un caso en biomedicina, podría suceder que, tras analizar miles de artículos sobre un patógeno, una IA podría sugerir una posible relación entre un mecanismo molecular y un fármaco existente originalmente desarrollado para otra enfermedad, señalando un nuevo rumbo experimental y actuando como una catalizadora de creatividad científica. De hecho, ya se reportan

casos donde la IA ha ayudado a predecir interacciones proteína-fármaco o a descubrir usos alternativos de medicamentos, hipótesis que luego son puestas a prueba en el laboratorio. Es importante señalar que estas hipótesis generadas por IA deben someterse al escrutinio científico habitual.

No menos importante es el apoyo de la IA en la organización del conocimiento. Herramientas de PLN pueden realizar *text mining* sobre miles de resúmenes o patentes, construyendo grafos de conocimiento que vinculan conceptos clave: por ejemplo, relacionando un gen con una ruta metabólica y con una enfermedad. Estos grafos semánticos asistidos por IA facilitan a los equipos de investigación ver “el bosque” en lugar de perderse entre los árboles de información aislada. En humanidades digitales, de manera análoga, se han empleado modelos de lenguaje para leer grandes colecciones de libros antiguos, extrayendo temas históricos comunes o influencias entre autores de distintas épocas, lo que sugiere nuevas hipótesis en historia cultural. En definitiva, la IA se está convirtiendo en un aliado intelectual del personal investigador: un lector incansable que resume, conecta y sugiere, ayudando a navegar la sobreabundancia de información y a iluminar caminos inexplorados para la investigación.

Diseño experimental y descubrimiento de nuevos materiales o fármacos

La IA abre horizontes totalmente nuevos en el descubrimiento científico, permitiendo diseñar experimentos y encontrar materiales o fármacos que antes hubiesen sido cuestión de azar o de largos procesos de ensayo y error.

En el diseño experimental, la IA está mejorando la planificación y ejecución. Técnicas de optimización bayesiana guiadas por IA ayudan a decidir qué experimentos realizar a continuación en proyectos de investigación. Si un químico tiene decenas de variables de reacción (temperatura, solvente, catalizador, etc.), una IA puede sugerir el siguiente conjunto de condiciones a probar que aporte más información o mejora, basándose en resultados previos. Esto acelera los ciclos experimentales y evita gastar recursos en pruebas poco prometedoras. Combinando robots de laboratorio con algoritmos de decisión, ya existen laboratorios autónomos que iteran experimentos de forma continua: formulan una hipótesis, realizan la prueba, analizan los datos con IA y vuelven a proponer la siguiente experimentación. Estos laboratorios auto-dirigidos han conseguido descubrir reacciones químicas y materiales novedosos con mínima intervención humana.

En ciencias de materiales, por ejemplo, los algoritmos de aprendizaje profundo están revolucionando la búsqueda de compuestos con propiedades específicas (superconductores, catalizadores, baterías más eficientes, etc.). Un caso notable es el desarrollado por DeepMind con técnicas de redes neuronales gráficas: su herramienta GNoME (Graph Networks for Materials Exploration) permitió predecir la estabilidad de millones de compuestos cristalinos potenciales. Como resultado, se descubrieron varios millones de nuevos cristales. Entre esos compuestos virtuales hay candidatos a superconductores, semiconductores avanzados y materiales para baterías de próxima generación, muchos de los cuales podrán ser sintetizados y probados experimentalmente en el futuro cercano.

En el ámbito de los fármacos y la biomedicina, la IA está logrando hitos igualmente impresionantes. Los modelos generativos (un tipo de IA capaz de crear nuevas muestras) han sido aplicados al diseño molecular de medicamentos. La farmacéutica Eli Lilly utilizó un sistema de inteligencia artificial para probar virtualmente millones de estructuras moleculares. Esta herramienta fue capaz de generar compuestos completamente nuevos, ausentes en sus bases de datos, lo que evidencia cómo la colaboración entre humanos y máquinas puede potenciar la creatividad en la investigación científica. Esta sinergia se ha comparado con tener un “colaborador artificial” que aporta ideas frescas.

Un avance que marcó un antes y un después en biología molecular es AlphaFold, el sistema de DeepMind para predicción de estructuras de proteínas. Tradicionalmente, determinar la forma 3D de una proteína era un proceso lento (por cristalografía o criomicroscopía) y miles de proteínas permanecían “sin resolver”. AlphaFold cambió ese panorama al demostrar que una IA entrenada con secuencias de aminoácidos y estructuras conocidas podía *aprender las reglas del plegamiento proteico*. Este logro, considerado uno de los más importantes de la IA en ciencia, permite comprender la estructura proteica a una escala jamás imaginada, y tiene implicaciones directas en el diseño de fármacos. Con las estructuras en mano, la comunidad científica puede identificar sitios activos donde pequeñas moléculas podrían unirse, acelerando el descubrimiento de medicamentos que actúen sobre proteínas. Técnicamente, AlphaFold emplea arquitecturas de aprendizaje profundo (especializadas en secuencias biológicas) para inferir estructuras a partir de secuencias. Su éxito demostró el poder de la IA para codificar el conocimiento científico (en este

caso, principios fisicoquímicos de plegamiento) y resolver problemas que desafiaban a los expertos durante décadas.

En síntesis, la IA está impulsando un nuevo paradigma en el descubrimiento científico, donde la generación de conocimiento – sea un material revolucionario o un fármaco salvador– ocurre más rápido y con enfoques antes inalcanzables, gracias a la creatividad aumentada y la búsqueda inteligente que aporta la inteligencia artificial.

Interpretación y visualización de resultados complejos

Con el aumento de la complejidad en los datos y modelos científicos, surge también la necesidad de interpretar y visualizar resultados. En este terreno, la IA juega dos papeles fundamentales: por un lado, ayuda a desentrañar las salidas de modelos complejos, y por otro, produce visualizaciones claras para comunicar hallazgos. Un desafío de los métodos de IA (especialmente los de aprendizaje profundo) es que a menudo actúan como “cajas negras”, generando predicciones acertadas pero difíciles de explicar. Esto ha motivado el desarrollo de técnicas de *IA explicable* (XAI, por sus siglas en inglés) que acompañan a los modelos en sus decisiones. Por ejemplo, existen algoritmos que analizan una red neuronal entrenada y resaltan qué variables o características de los datos influyen más en su resultado. En el caso de imágenes, un método llamado Grad-CAM puede mostrar un mapa de calor sobre la imagen indicando las regiones que la red neuronal considera más relevantes para, por ejemplo, clasificar un tumor en una radiografía. Así, el médico no solo recibe la predicción “*positivo/negativo*”, sino también una visualización que le indica *dónde* está mirando la IA en

la imagen, lo que aporta confianza y posibilidad de verificación. Estas herramientas de interpretabilidad son cruciales para integrar la IA en la práctica científica, pues proporcionan transparencia: permiten al experto entender (al menos parcialmente) la base de una recomendación algorítmica, facilitando la validación independiente. Además, la IA ayuda a la visualización: por ejemplo, algoritmos de reducción de dimensionalidad (como *t-SNE* o *UMAP*, basados en aprendizaje no supervisado) proyectan datos de alta dimensión a gráficos 2D o 3D. Esto es útil para visualizar la relación entre miles de genes a la vez en estudios genómicos o para mapear cómo se agrupan galaxias en cosmología. La IA, por tanto, actúa como intérprete entre el mundo de los números y ecuaciones de alta complejidad y la intuición humana, produciendo representaciones comprensibles de fenómenos complejos.

En cuanto a la interpretación de resultados empíricos, un caso muy relevante es el de la medicina. Los sistemas de IA aplicados al diagnóstico no solo emiten un dictamen, sino que complementan al especialista resaltando patrones sutiles que este podría pasar por alto. En radiología, la IA actúa como un segundo par de ojos infatigable que señala las regiones potencialmente patológicas en cada imagen, facilitando la interpretación médica. Esto no solo aumenta la sensibilidad diagnóstica, sino que mejora la eficiencia al priorizar la atención del especialista donde realmente importa. Al superponer las indicaciones de la IA sobre la imagen (por ejemplo, rodeando un nódulo difícil de ver), el médico obtiene una visualización clara de por qué la máquina considera sospechosa esa mamografía. Este enfoque de colaboración ya se extiende a otras áreas: en patología digital, sistemas de visión por computador escanean láminas histológicas completas y marcan con colores

células tumorales aisladas; en neurología, algoritmos que analizan electroencefalogramas destacan los momentos donde se produce una convulsión epiléptica; en economía, modelos predictivos indican qué variables (inflación, desempleo, etc.) están contribuyendo más a un cambio observado en la proyección financiera, ayudando a los analistas a interpretar las causas de una tendencia. La clave es que la IA puede triangular información compleja y presentarla de forma accesible.

Redacción, traducción y publicación de resultados científicos

La faceta comunicativa de la ciencia también puede beneficiarse del uso de la inteligencia artificial. Tradicionalmente, escribir un artículo científico en inglés (idioma dominante de la ciencia) podía ser un obstáculo para investigadores(as) no nativos, o consumir un tiempo considerable en revisiones de estilo y corrección gramatical. Hoy en día, existen asistentes de escritura impulsados por IA que apoyan desde la gramática hasta la claridad expositiva. Herramientas como Grammarly o ProWritingAid, basadas en ML, revisan el texto científico señalando errores, frases confusas, repeticiones e incluso verificando la consistencia terminológica. Estas herramientas pueden integrarse en los procesadores de texto y actuar casi en tiempo real mientras se escribe, sugiriendo mejoras. De forma similar, la traducción automática ha alcanzado un nivel de calidad utilizable en contextos académicos: servicios como DeepL o Google Translate (potenciados por redes neuronales de traducción) permiten traducir un manuscrito completo del español al inglés en segundos, preservando en gran medida el sentido técnico. Aunque un humano debe luego pulir detalles, esta asistencia acelera enormemente la preparación de manuscritos

multilingües.

Las herramientas de IA permiten generar textos completos a partir de indicaciones, por ejemplo, se puede indicar al modelo: "*resume los resultados de este estudio en un párrafo*" o "*sugiere maneras de explicar tal concepto para una audiencia general*". El modelo entonces produce un texto que se puede adoptar o modificar. Esto resulta especialmente útil en etapas iniciales de escritura o cuando se quiere adaptar un contenido a diferentes formatos (por ejemplo, convertir un trabajo denso en un texto divulgativo). Herramientas experimentales como NotebookLM permiten cargar un conjunto de documentos (apuntes, artículos) y hacer preguntas al respecto, recibiendo respuestas elaboradas con citas exactas de los documentos. En el proceso de publicación científica, la IA también está presente: algunas revistas emplean software de IA para revisar automáticamente manuscritos en busca de plagio, detección de manipulación de imágenes, o incluso para sugerir posibles revisores expertos en función del contenido del artículo. Es decir, la IA asiste no solo a los autores sino también a editores y revisores, haciendo más eficiente el ciclo de evaluación por pares.

Un aspecto interesante es la traducción y adaptación cultural de los resultados científicos. Un equipo de investigación internacional puede apoyarse en IA para generar versiones del mismo artículo en varios idiomas, o para elaborar resúmenes ejecutivos dirigidos a distintos públicos (por ejemplo, a patrocinadores gubernamentales o a medios de comunicación generalistas). En humanidades digitales, proyectos de edición de textos clásicos utilizan IA para sugerir traducciones anotadas y acelerar la publicación bilingüe de manuscritos históricos, algo que antes podía llevar años de labor

humana.

Es importante destacar que, si bien la IA facilita la redacción, no sustituye el rigor que el personal investigador debe aportar. Las herramientas de generación de texto pueden producir con facilidad afirmaciones incorrectas o referencias inventadas (un fenómeno conocido como *alucinaciones* de los LLM). Por ello, su uso requiere siempre revisión y verificación. De hecho, muchos editores científicos están estableciendo pautas éticas: por ejemplo, algunas revistas solicitan que se declare el uso de IA en la preparación del artículo, y han dejado claro que no aceptarán modelos de IA como coautores. En resumen, la IA se ha incorporado también en la fase final del quehacer científico, pulido y diseminación del conocimiento, rompiendo barreras idiomáticas y estilísticas, y prometiendo una ciencia más accesible y colaborativa globalmente.

La importancia de los prompts

Un prompt es un texto o instrucción que se utiliza para que una herramienta basada en un LLMs proporcione una respuesta. El prompt puede incluir desde una pregunta sencilla hasta indicaciones complejas con contexto adicional, ejemplos y formato esperado. Lógicamente, la calidad de las frases o preguntas es un factor de gran importancia para lograr una conversación exitosa con la herramienta. En esencia, un buen prompt aprovecha mejor las

capacidades del modelo, mientras que uno deficiente las desaprovecha.

Técnicamente, un prompt es el *input* (entrada de texto) que se le da a un modelo de lenguaje para indicarle qué tarea realizar. Formalmente, podría consistir en una pregunta, una instrucción, un conjunto de datos o una combinación de estos elementos. Por ejemplo, si se indica a un modelo “*Resume el siguiente párrafo en una frase*”, todo ese enunciado funciona como prompt indicando al modelo que genere un resumen.

La diferencia de resultados entre un prompt genérico y otro específico puede ser sorprendente. En contextos de investigación científica, donde la precisión es clave, invertir esfuerzo en redactar un buen prompt puede determinar si la respuesta de la IA es aprovechable. En resumen, un prompt bien redactado maximiza la utilidad de la IA, obteniendo respuestas más precisas, completas y alineadas con lo que se precisa.

Principios para diseñar prompts efectivos

A la hora de elaborar prompts de calidad, existen principios que sirven de guía. Entre los más importantes están el *contexto*, la *claridad*, la *especificidad*, y el *lenguaje objetivo*. Estos criterios ayudan a estructurar la petición de forma que el modelo entienda exactamente qué se espera.

Proporcionar contexto en el prompt es ofrecer al modelo la información de fondo o la situación en la que debe basarse para

generar su respuesta. El contexto puede incluir datos previos, la motivación de la pregunta o incluso asumir un rol. Por ejemplo: “*Eres un doctorando en Química. Basándote en los datos experimentales [X], explica qué hipótesis podrían derivarse*”. El escenario y el rol proporcionado al modelo antes de la pregunta, le ayuda a formular una respuesta más pertinente. Un prompt contextualizado reduce la ambigüedad y aumenta la coherencia de la respuesta con la situación deseada.

La claridad implica que el prompt esté formulado en un lenguaje sencillo y directo, sin frases enrevesadas ni términos confusos. Hay que evitar modismos locales, jerga excesiva o formulaciones que puedan interpretarse de más de una manera. Un prompt claro establece sin rodeos la tarea: por ejemplo, en lugar de decir “*Me pregunto si podrías proporcionarme alguna información sobre bases de datos científicas especializadas en Psicología*”, una versión clara sería “*Elabora un listado de bases de datos científicas especializadas en Psicología indicando una característica principal de cada una*”. La claridad también se refiere a no mezclar varias solicitudes en una sola frase y a usar una gramática y ortografía correctas.

Ser específico significa brindar detalles concretos sobre lo que se desea. Un prompt específico acota el tema y el formato de la respuesta esperada. Por ejemplo, si se necesita que la IA ayude a obtener datos sobre un fenómeno, no es lo mismo pedir “*Háblame del cambio climático*” que “*Resume en 3 párrafos los efectos del cambio climático en los patrones de lluvia en España citando en APA al menos un artículo científico*”. En el segundo prompt se indican claramente la extensión (3 párrafos), el enfoque geográfico

(España), el aspecto a cubrir (patrones de lluvia) y hasta el estilo (incluir referencia en APA a un artículo científico). Estos detalles específicos guían al modelo para que su respuesta sea mucho más útil. La especificidad evita respuestas genéricas; al contrario, orienta a la IA hacia la información o formato exacto que se requiere.

El lenguaje objetivo se refiere a especificar en el prompt el idioma, estilo o tono en que se solicita la respuesta. Los LLM pueden generar texto en múltiples idiomas y registros, por ejemplo, la comunicación científica. Indicar el público objetivo en el prompt es importante: ¿es un resumen para expertos, para divulgación general, o para estudiantes? Asimismo, hay que precisar si la respuesta se necesita en un formato específico (por ejemplo, una lista numerada, una tabla, etc.). Este tipo de indicaciones asegura que el modelo adapte su salida al formato y tono que mejor se ajuste a nuestras necesidades.

Tipos de prompts

Existen diferentes tipos de prompts que pueden emplearse al interactuar con un LLM, dependiendo de la naturaleza de la pregunta y del tipo de respuesta esperada. A continuación, se describen varios tipos, incluyendo *prompts abiertos o cerrados, secuenciales, multitarea e instructivos*, entre otros.

Los prompts abiertos son aquellos que no restringen la respuesta a un formato o contenido específico, sino que dejan libertad al

modelo para generar la respuesta. Suelen fomentar respuestas amplias y creativas porque no hay una única respuesta "correcta". Por ejemplo: "*Describe posibles aplicaciones futuras de la inteligencia artificial en la educación superior*". Esta pregunta abierta invita al modelo a explorar múltiples ideas. Los prompts abiertos son útiles para brainstorming, para obtener explicaciones extensas o puntos de vista variados. Por su parte, los prompts cerrados buscan respuestas concretas y específicas, a menudo delimitando opciones. Un prompt cerrado típicamente espera una respuesta breve o incluso una opción binaria (sí/no, verdadero/falso) o un dato puntual. Por ejemplo: "*¿En qué año fue inventado el autogiro?*" es un prompt cerrado que apunta a una respuesta específica (1923, en este caso). Este tipo de prompt es útil cuando se busca información puntual o una confirmación.

Los prompts secuenciales son una técnica que consiste en encadenar varias interacciones o estructurar una serie de pasos dentro de una misma solicitud con el fin de desglosar una tarea compleja. En lugar de pedir todo de una vez, se guía al modelo paso a paso. Esto se puede hacer de dos maneras: mediante múltiples turnos de conversación (hacer una pregunta inicial, luego basándonos en la respuesta, hacer la siguiente, y así sucesivamente) o dentro de un solo prompt dividiendo la instrucción en partes. Por ejemplo, se podría iniciar con "*Dame una lista de los 10 factores más importantes que contribuyen al crecimiento económico.*" Una vez obtenida la lista, seguir con "*Ahora, elabora un párrafo explicando cómo cada factor influye en la economía*". Los prompts secuenciales proporcionan mayor

control sobre cada etapa del proceso y son ideales para tareas largas o complejas que requieren estructura.

Un prompt multitarea es aquel que solicita varias acciones o tipos de respuesta en una misma interacción. Por ejemplo: *"Resume el siguiente informe y elabora dos mensajes para difundirlo en la red social X, manteniendo un tono formal.* Estos prompts multitarea pueden ahorrar tiempo al obtener una respuesta comprensiva de una sola vez. Sin embargo, también conllevan el riesgo de que el modelo responda solo parcialmente si alguna instrucción se " pierde" entre las demás. Para usarlos eficazmente, es importante estructurar muy bien el pedido, listando claramente cada sub-tarea o usando delimitadores para separarlas. Por ejemplo, especificar: *"Tarea 1: ...; Tarea 2: ..."*. Aun así, en ocasiones es preferible dividir en prompts secuenciales en vez de un único prompt multitarea.

Los prompts instructivos o *prompts de instrucción* son aquellos formulados como órdenes o indicaciones directas al modelo para realizar una tarea específica. En lugar de hacer una pregunta abierta, se indica al modelo qué debe hacer. Ejemplos de prompts instructivos son: *"Genera un resumen ejecutivo para el siguiente informe"*, *"Traduce al español el texto proporcionado"*, *"Escribe un código en Python que calcule la media de una lista de números"*. Este tipo de prompt suele comenzar con un verbo en modo imperativo que indica una acción a realizar. Los prompts instructivos son muy efectivos cuando se conoce con precisión la salida esperada (formato y contenido). También son útiles para establecer roles: por ejemplo, *"Actúa como un experto en bibliometría y evalúa los indicios de calidad de este artículo"*. En

general, para la mayoría de las tareas concretas usar un estilo instructivo es recomendable porque reduce la ambigüedad.

Buenas prácticas

Al redactar prompts conviene seguir buenas prácticas específicas de formato y técnica que ayudan a evitar malentendidos y a guiar a la IA de manera óptima. A continuación, se destacan algunas:

- Usar formatos estructurados (Markdown, listas, etc.) en la respuesta deseada: si se precisa que la respuesta tenga una cierta estructura (por ejemplo, una lista con viñetas, una tabla, encabezados Markdown, etc.), es útil indicarlo en el prompt. Se puede escribir algo como: *"Proporciona la respuesta en formato de lista con viñetas"* o *"Muestra los resultados en una tabla Markdown con columnas X e Y"*. Los LLM son capaces de formatear la salida si se les pide claramente. Esto es muy valioso, por ejemplo, al solicitar un resumen tabulado de datos experimentales o una lista de referencias bibliográficas formateadas. Una respuesta estructurada no solo es más legible, sino que facilita su uso en documentos, plataformas que soporten Markdown o análisis posteriores.
- Delimitar claramente secciones o contexto dentro del prompt: Cuando se incluye contenido largo o varias partes en el prompt (por ejemplo, un bloque de texto que el modelo debe analizar, seguido de una pregunta), es importante separar estas partes de forma inequívoca. Una

técnica común es usar *delimitadores* como triple comillas "''' o backticks ` ` ` para encerrar el texto proporcionado, o etiquetas como “*Texto:*” seguido del contenido y luego “*Instrucción:*” antes de la pregunta. Esto evita que el modelo se confunda sobre qué forma parte de la pregunta y qué es dato proporcionado. Por ejemplo:

Analiza el siguiente texto y luego responde a la pregunta indicada.

Texto: "'''

[Aquí iría un párrafo extraído de un artículo científico]

'''

Instrucción: ¿Cuál es la hipótesis principal del autor?

En este prompt, mediante el uso de "''' se delimita claramente el texto a analizar. OpenAI recomienda separar la instrucción del contexto con símbolos especiales para mayor claridad

- Incluir ejemplos en el prompt (few-shot prompting): Proporcionar ejemplos concretos de lo que se busca puede mejorar significativamente la calidad de la respuesta. Esta técnica consiste en mostrar al modelo una pareja de entrada-salida de ejemplo antes de hacer nuestra solicitud real. Por ejemplo, si se considera útil que ChatGPT genere metáforas científicas, puede antecederse la solicitud con una indicación específica en ese sentido: “*Ejemplo: ‘El ADN es como una biblioteca donde cada gen es un libro de instrucciones.’ Ahora crea una metáfora similar que explique el funcionamiento de una célula.*”. Los ejemplos actúan como guía del estilo o formato deseado.

- Contraejemplos y órdenes negativas: Así como los ejemplos positivos ayudan, a veces es útil aclarar la respuesta no deseada. Se pueden incluir contraejemplos u órdenes explícitas de exclusión. Por ejemplo: “*No incluyas definiciones triviales, enfoca la respuesta en aspectos novedosos*”. Un uso inteligente de contraejemplo es mostrar un formato o estilo incorrecto y señalar que se evite: “*Pregunta: [...] Respuesta incorrecta: (un texto). Ahora, brinda una respuesta correcta y bien fundamentada*”. Esto puede recalibrar al modelo si tiende a irse por un camino no deseado. En general, es más seguro indicar directamente prohibiciones o preferencias (por ejemplo: “*no menciones valores de indicadores normalizados por debajo de 1*”, que dar una respuesta errónea completa. Las órdenes negativas funcionan como límites que el modelo normalmente respetará.
- Iteración y refinamiento: Por último, una buena práctica es estar dispuesto a iterar el prompt. Si la primera respuesta no es satisfactoria, conviene refinar la solicitud y volver a intentarlo. Esto no siempre implica empezar de cero; a veces se puede aprovechar la misma conversación. Por ejemplo, usando indicaciones de seguimiento: “*Por favor, profundiza más en tal aspecto*” o “*Reformula la respuesta anterior en un tono más académico*”. O bien, editar el prompt original para añadir esa aclaración que faltó. La interacción con un LLM es un proceso dinámico de diálogo.

Siguiendo estas buenas prácticas, se incrementan las probabilidades de obtener de la IA respuestas precisas, útiles y

formateadas a nuestra medida, reduciendo a la vez malinterpretaciones. Esto es especialmente importante cuando se aplica ChatGPT u otros LLM a tareas de investigación.

A continuación, se presenta un ejemplo de prompt:

[Rol]

Actúa como un *experto en estilos de citación académica*, con profundo dominio de normas como APA, MLA, Chicago, Vancouver e IEEE.

[Contexto]

Estás asistiendo a estudiantes, investigadores o profesionales que necesitan presentar sus referencias bibliográficas de forma impecable en trabajos académicos.

[Tarea]

Tu responsabilidad es revisar y formatear correctamente listas de referencias bibliográficas en el estilo que se te indique. Debes aplicar con precisión las normas del estilo solicitado, incluyendo:

- Orden y formato de los elementos
- Puntuación exacta
- Uso correcto de cursivas y mayúsculas
- Adaptación de los nombres de los autores
- Inclusión de DOI o URL cuando sea necesario
- Orden alfabético o numérico, según el estilo

[Ejemplo de entrada]

Formatea en **estilo APA (7^a edición)** las siguientes referencias que están sin formato:

- Beck, Aaron T. *Terapia cognitiva de la depresión*. Paidós. 1983.

[Formato esperado de salida]

Beck, A. T. (1983). *Terapia cognitiva de la depresión*. Paidós.

[Tono]

Utiliza un tono formal, claro y profesional. La salida debe estar lista para copiar y pegar en un trabajo académico.

Otros ejemplos de buenas prácticas

Los avances en inteligencia artificial han permitido la creación de **asistentes o agentes personalizados** que facilitan diversas tareas investigadoras. Una de las herramientas más destacadas en este ámbito es la posibilidad de crear agentes de inteligencia artificial dentro de la plataforma ChatGPT, conocidos como GPTs personalizados. Estos asistentes pueden configurarse fácilmente para cumplir funciones específicas, lo que los convierte en asistentes virtuales altamente adaptables a las necesidades de distintos proyectos de investigación.

Un GPT personalizado consiste en una versión de ChatGPT cuya conducta puede ser modificada mediante instrucciones detalladas, con el fin de orientarlo hacia un propósito determinado. Por ejemplo, se puede crear un agente que actúe como asistente de investigación en temas de ciencias sociales, que emplee un lenguaje técnico adecuado al nivel académico, que sea capaz de resumir documentos, proponer hipótesis o formatear referencias bibliográficas. Además, es posible cargar archivos y proporcionar referencias específicas que el agente puede utilizar como contexto, lo que mejora significativamente la relevancia y precisión de sus respuestas.

Una de las ventajas más significativas de estos agentes es que pueden compartirse fácilmente con otras personas mediante un enlace. Esto permite que colegas, estudiantes u otros investigadores puedan interactuar con el agente desde sus propias cuentas gratuitas de ChatGPT, sin necesidad de una suscripción de pago. De esta manera, se fomenta la colaboración y el acceso equitativo a herramientas avanzadas de apoyo a la investigación.

La Biblioteca de la UNED ha diseñado varios asistentes de apoyo a la investigación. A continuación, se presenta un ejemplo de un prompt diseñado para un agente de ChatGPT.



Actúa como un experto en normativas de citación académica, con profundo conocimiento de estilos como APA, MLA, Chicago, Vancouver, IEEE, entre otros.

Heurísticas a seguir:

Prioriza la precisión tipográfica: Asegúrate de que todos los elementos estén puntuados y formateados exactamente según el estilo especificado (cursivas, comillas, paréntesis, etc.).

Autocompleta solo si es seguro: Si una referencia está incompleta, intenta completarla solo si puedes confirmar la fuente en bases confiables como Crossref, Google Scholar o el sitio editorial. De lo contrario, usa un marcador como: [falta el año], [falta el DOI].

Abrevia títulos de revistas si el estilo lo requiere: Por ejemplo, en estilos como Vancouver o IEEE, utiliza abreviaciones oficiales de las revistas (según listas como NLM Catalog o ISO 4). Si no estás seguro, mantén el título completo y marca con [verificar abreviatura].

Tu tarea es formatear correctamente un listado de referencias bibliográficas según el estilo de citación que se especifique en la consulta. Aplica con precisión las normas del estilo, incluyendo:

Puntuación y formato (cursivas, uso de comillas, etc.)

Orden de los elementos: autor, año, título, fuente, etc.

Orden alfabético o numérico según el estilo. Alfabéticamente por autor (APA, MLA, Chicago...) Numéricamente (Vancouver, IEEE...)

Inclusión de DOI o enlaces cuando sea requerido

Adapta correctamente los nombres de autores iniciales y formato (apellido, inicial; orden, etc.), así como los años de publicación.

Abrevia el título de las revistas en caso de que el estilo lo exija.

Salida esperada:

Una lista de referencias formateadas y ordenadas según el estilo indicado, lista para copiar y pegar en un trabajo académico.

Si alguna referencia está incompleta, marca claramente el dato faltante con: [falta el año], [falta el DOI], etc.

[rol] [reglas] [tarea] [formato]

La principal diferencia del prompt del asistente con el “prompt de usuario” radica en la incorporación de reglas o estrategias para un mejor funcionamiento.

El profesor Antonio Julio López Galisteo, de la Universidad Rey Juan Carlos (URJC), ha desarrollado [GPT master maker](#), un asistente de inteligencia artificial que ayuda a mejorar y perfeccionar las instrucciones de los asistentes que se deseen diseñar. Por su parte, el profesor Oriol Borras (URJC) es el responsable del sitio web iaorioltic.wikidot.com dedicado a la recopilación y difusión de recursos relacionados con la inteligencia artificial (IA) aplicados a la educación y a la investigación. Se trata de un recurso especialmente útil para el personal docente e investigador interesado en integrar la IA en sus prácticas educativas y proyectos de investigación.

Por su parte, la Red de Bibliotecas Universitarias y Científicas Españolas (REBIUN) ha creado un [Observatorio de Inteligencia Artificial](#) que tiene como objetivo conocer, analizar y evaluar las principales iniciativas en materia de inteligencia artificial para determinar su incidencia y aplicabilidad en las bibliotecas universitarias y científicas. Entre sus recursos, ofrece guías,

informes y materiales de formación que destacan por su utilidad para el sector universitario, incluyendo herramientas de IA de apoyo a la docencia y a la investigación, así como información sobre el uso legal y ético de la inteligencia artificial.

The screenshot shows the homepage of the REBIUN Observatory of Artificial Intelligence. At the top, there's a navigation bar with links to INICIO, RECURSOS, PROYECTOS, FORMACIÓN, and USO LEGAL Y ÉTICO. Below the navigation, there's a section titled "Herramientas de apoyo a la investigación" (Tools for research support) which displays three cards: OSF, CONSENSUS, and SCOPUS AI.

- OSF:** Described as a tool for managing research projects, it's a free, open-source platform for storing and sharing documents.
- CONSENSUS:** Described as a search engine for research, it allows users to search across different databases and offers bibliometric information.
- SCOPUS AI:** Described as a tool for generating brief summaries of documents indexed in the Scopus database, based on instructions or indications.

Uso ético de la inteligencia artificial generativa

Pese a sus enormes ventajas, la integración de la IA en la investigación científica plantea desafíos éticos y metodológicos que deben ser abordados con mirada crítica. Uno de los temas más discutidos es el problema de la “caja negra”: muchos algoritmos de

IA, especialmente los basados en redes neuronales profundas, no ofrecen explicaciones claras de cómo llegan a sus conclusiones. Esta opacidad es preocupante en ciencia, ya que el método científico exige transparencia y reproducibilidad. Si un sistema de IA propone una hipótesis o identifica un patrón en los datos, el personal investigador necesita poder seguir la pista de por qué lo hizo para confiar en el hallazgo. La falta de interpretabilidad puede llevar a sesgos ocultos: el modelo podría estar tomando en cuenta correlaciones espurias o prejuicios presentes en los datos de entrenamiento, y sin una ventana a su razonamiento interno, estos sesgos pasarían inadvertidos.

Otro desafío ético es el de la validez y la veracidad de la información generada por IA. Los modelos generativos de lenguaje pueden producir texto muy convincente, pero no garantizan la verdad factual. En el contexto de la revisión de literatura, esto puede ser peligroso: si un investigador(a) se apoyara ciegamente en un resumen hecho por IA, podría incorporar afirmaciones inexistentes en su trabajo. Por ello, se insiste en que la IA sea una herramienta de apoyo; las salidas de la IA tienen que ser siempre verificadas con las fuentes originales. Asimismo, existe el riesgo de la fabricación involuntaria de datos o análisis. Un modelo podría “rellenar” lagunas en datos experimentales con estimaciones que parezcan plausibles, pero que no se basan en la realidad, lo que comprometería la integridad científica si no se detecta. Desde una perspectiva de integridad académica, también se discute si el uso de IA para escribir textos puede rozar el plagio o la autoría indebida. Aunque la IA genera frases nuevas, podría hacerlo

calcando demasiado de sus datos de entrenamiento (por ejemplo, reproduciendo trozos de artículos previos sin citarlos).

La privacidad y la confidencialidad son otras consideraciones: en campos como la medicina o las ciencias sociales, entrenar IA implica usar datos de personas (historias clínicas, encuestas, etc.). Es imperativo asegurarse de que esos datos estén anonimizados y almacenados con seguridad, y que los modelos resultantes no puedan filtrar información sensible. Por ejemplo, un modelo de lenguaje entrenado con textos clínicos podría, si no se cuida, llegar a regenerar fragmentos reconocibles de un caso real. Las regulaciones imponen fuertes restricciones al manejo de datos personales que las y los científicos deben cumplir.

En este sentido, el marco normativo europeo aporta una guía clara para una IA responsable. El Reglamento de Inteligencia Artificial de la Unión Europea (AI Act), aprobado en 2024, constituye la primera legislación integral sobre IA a nivel global. Clasifica los sistemas según su nivel de riesgo y establece requisitos proporcionales: desde sistemas de bajo riesgo hasta aquellos de alto riesgo, como los aplicados en salud o evaluación académica, que exigen trazabilidad, supervisión humana y transparencia. La investigación científica que utilice IA debe alinearse con estas exigencias, adoptando prácticas que aseguren la protección de los derechos fundamentales.

Además, el AI Act refuerza la necesidad de cumplir con el Reglamento General de Protección de Datos (RGPD), reiterando la importancia de la privacidad desde el diseño. Esto obliga a que los

datos utilizados para entrenar sistemas de IA estén adecuadamente anonimizados y protegidos, minimizando cualquier posibilidad de reidentificación o filtración. Así, se refuerza el compromiso con una ciencia abierta, pero también ética y segura.

En cuanto al desafío de la gobernanza hay que destacar la adopción de la Carta Iberoamericana de Inteligencia Artificial en la Administración Pública, aprobada en 2022 por el Centro Latinoamericano de Administración para el Desarrollo (CLAD). Esta carta proporciona un marco de referencia para orientar a los gobiernos en el uso de la IA dentro del sector público. Propone principios clave como el respeto a los derechos humanos, la transparencia, la rendición de cuentas, la seguridad de los datos, la inclusión digital y la sostenibilidad ambiental. Además, la Carta enfatiza la necesidad de que las administraciones públicas desarrollen marcos normativos internos, establezcan comités éticos, implementen mecanismos de auditoría algorítmica y promuevan la participación ciudadana en la formulación de políticas relacionadas con la IA. También impulsa la cooperación regional entre los países iberoamericanos para intercambiar buenas prácticas, fortalecer capacidades institucionales y avanzar hacia una gobernanza común basada en valores compartidos.

Por último, está el desafío de no desplazar el pensamiento crítico humano. La IA tiende a dar respuestas rápidas y definitivas, pero la ciencia prospera en el cuestionamiento y la duda. Existe el riesgo de que investigadores(as) jóvenes, habituados a consultar siempre a un asistente de IA, pierdan habilidad para resolver problemas por sí mismos o para discernir sutilezas que la IA pasa por alto. La

integración ética y crítica de la IA implica usarla como extensión de nuestras capacidades, no como sustituto de estas. Los científicos deben seguir formulando preguntas creativas, desconfiando sanamente de las respuestas fáciles y aplicando su experticia para supervisar a la máquina.

En palabras simples: la IA puede calcular y correlacionar, pero el sentido y el contexto lo aportan las personas. Un enfoque verdaderamente reflexivo y ético requiere entrenar no solo modelos, sino también personas usuarias responsables. En este nuevo escenario, donde tecnología y conocimiento se entrelazan, el mayor reto no será lo que la IA pueda hacer, sino lo que nosotros decidamos hacer con ella. Solo así será posible una ciencia que no solo avance, sino que también se sostenga sobre valores sólidos de integridad, equidad y responsabilidad.